



Australian
Bureau of
Statistics

Responsible Access to ABS

Confidentialised Unit Record Files (CURFs)



Training Manual



Responsible Access to ABS Confidentialised Unit Record Files (CURFs) Training Manual

Edition 2, March 2005

© Commonwealth of Australia 2005

Produced by the Australian Bureau of Statistics

1. About this manual

This manual tells users of ABS CURFs about their role in protecting the confidentiality of unit record data. Features of the manual are

- Main points at the top summarise major issues
- Key user responsibilities are highlighted
- Details and examples fill the remainder of each page.



Your Responsibility

Prior to using ABS CURFs users should read this training manual so that they clearly understand their obligations in accessing the data under the legal undertaking they sign with the ABS.

Legislation allows the Australian Statistician to release unit data provided that this is done:

“ in a manner that is not likely to enable identification of a particular person or organisation to which it relates ”

The ABS has adopted a **manner of release** that protects the confidentiality of ABS respondents in three ways:

- Control the **detail** available on the confidentialised unit record file (CURF)
- Provide **modes of access** appropriate to the level of detail available
- Place restrictions on how the data is used, as set out in a legal **undertaking** to be signed by each user and their organisation

The **restrictions on user behaviour** are critical to the ABS being able to release unit data at the levels of detail that it does. By requiring users to read this manual, the ABS helps guarantee that users understand their responsibilities and take them seriously.

To complete the "Responsible Access to ABS CURFs" training you must read carefully the **main points** and summary of **your responsibilities** that appear at the top of each page. The more detailed material on each page is provided to help you understand these main points. Some of this material is illustrative rather than exhaustive, and if you need more guidance on any issues you can contact the ABS (see section 10 for contact details).

For each **mode of access** that you use you should read the full detail in the corresponding section of this manual. The ABS may provide data in one of three modes of access:

- CD-ROM
- Remote Access Data Laboratory (RADL)
- ABS Site Data Laboratory (ABSDL).

2. Why protect identity?

CURFs contain information provided by individuals or organisations. The ABS is careful not to disclose the identity of these respondents - to do so would not only breach legislation but also breach the trust the respondent places in the ABS. This trust is a key factor in maintaining the high quality of ABS data.



Your Responsibility

As a user of ABS unit record files, you must not seek to identify the individuals or organisations represented on the files.

The ABS is highly respected within the Australian community for its standards of **impartiality** and **confidentiality**. This means that people can feel confident in providing personal information, or information about their organisation, to the ABS.

Public confidence in the ABS would diminish if the identity of the individuals or organisations on a unit record file were disclosed to individuals outside the ABS. This would not only be a breach of trust with the survey respondents, but also a breach of the legislation under which the information was collected. Such a breach could harm the reputation of the ABS, and lead to a reduction in the quality of the survey estimates.

ABS releases unit data under conditions designed to prevent this scenario arising. Unfortunately, simply removing name and address details from files does not completely prevent the identification of survey respondents. Key steps the ABS takes are:

- Reducing detail on the CURFs, or modifying values, to make it difficult for units to be identified;
- Placing constraints on how the user can access the data and on what they do with it;
- Requiring anyone who uses the data, and the organisations that they represent, to sign an undertaking with the ABS that identifies the legal requirements to use the data; and
- Obtaining user commitment to confidentiality principles.

A fundamental constraint on users is that they must not seek to identify, by any means whatsoever, the individuals or organisations represented on the CURFs. Other constraints limit behaviour, such as matching to other unit datasets (see section 4), that could increase the likelihood of identification.

3. Spontaneous recognition

Identification could arise if a user recognises unusual characteristics of a person or household on the CURF as corresponding to those of someone they know of. The ABS reduces the likelihood of such "spontaneous recognition" by limiting the detail available on CURFs and by modifying some items for unusual units.



Your Responsibility

If you see ABS unit record information that appears to correspond to some individual or organisation you know, do not attempt to confirm the identity of the unit.

An individual can be identified even though name and address are not released. Spontaneous recognition occurs when a user sees unusual characteristics of a unit that match someone they know, and proceeds to check this identification based on other attributes of the unit. In outline, spontaneous recognition proceeds as follows:

User notes unusual combination on file:
Sex=female, country of birth=Turkey, occupation=doctor

User recognises they know such a person:
"That could be Mary"

User checks other characteristics:
Yes, Mary is 43 years old and arrived in 1995

Result: person is identified
Worst case scenario: User finds out sensitive information that Mary provided confidentially.

How "spontaneous recognition" happens

The ABS minimises the risk of spontaneous recognition by limiting detail on the CURFs. In addition, a small number of particularly unusual records may have some of their values changed. Thus it is quite likely that any apparent recognition of an individual or organisation will be incorrect. You should not attempt to confirm the identity of any unit record, nor should you inform anybody else of a potential recognition.

4. Matching to lists

Some items on a unit record file may correspond to items available on another list of units, perhaps an administrative dataset. Matching the two datasets on these items could lead to identification of the units.



Your Responsibility As a user of an ABS CURF, you must not match the unit data to any other unit information.

The undertaking that users sign to gain access to ABS unit data specifically **disallows matching** the information with any other unit level list. The prohibition against matching also applies to very fine tabulations in which some cells have only a single contributor, since these cells constitute unit level information.

Even if the matching has a valid statistical purpose, it is still **prohibited** in the undertaking you have signed. This prohibition applies to any linking or merging of unit files, such as "statistical matching". For example, matching includes attempts to add extra items to the ABS unit data files by attaching values derived from similar units in another collection. If you wish to extend the file in this way, you must contact ABS and describe the process you wish to carry out and the purpose of it (see p.10 for contact information). ABS will assess whether this constitutes matching that might be permitted, subject to an agreed protocol and the signing of a separate undertaking.

The **RADL and ABSQL modes of access** are designed to preclude the possibility of large-scale matching by keeping the unit record data within the ABS and monitoring the use that is made of the data. These methods provide for users to see enough unit records to support the checking of their code and analysis, while preventing large-scale matching.

A particularly high matching risk arises in a file of persons in which the **household or family** that each person belongs to is identified. For example,

- Households can be classified into quite fine categories e.g. 1 adult with 2 children.
- Such a multiple person household could then be further classified by other characteristics such as the age and sex of the three individuals.
- Applying such a classification to an administrative list may reveal some categories that contain only one household in the population.
- Any households on the unit record file that fall into those categories would then potentially be identified as the same ones as on the administrative list.

The ABS is particularly careful to limit the detail available for items that may appear on administrative lists, to limit the risk of identification from matching.

5. Keeping the unit information secure

Access to unit information is restricted to individuals who have signed an undertaking with the ABS for access to the data.



Your Responsibility

You must keep all unit data secure so that it is not accessible by anyone who does not also have authorised access to the unit record file. Results from your analyses that may reveal unit information should also be kept secure.

You must keep ABS unit information where it cannot be accessed by anyone except those specifically permitted by the ABS (i.e. researchers within your organisation who have been approved by the ABS to access this unit data after signing the required undertaking).

The following are guidelines for storing the unit information and results of your analyses.

- Any **computer** on which ABS unit information is stored must be kept in a locked room and/or have password protection to prevent access by others.
- Any information or results stored on a **computer network** must be kept in a directory with suitably restricted access.
- Any **printouts** or any **physical medium** containing unit information (such as a CD-ROM provided by the ABS) must be kept in a locked room or secured in a locked cabinet when the researcher is not present.
- You must keep track of datasets and printouts that reveal unit data, so that they can be **destroyed** when they are no longer needed. Printouts can be destroyed securely by shredding. Your organisation must also provide for a secure way to delete computer files so that they cannot be recovered by an unauthorised person.

These restrictions apply to unit data and to outputs from your analysis that reveal information for individual units.

- Examples of data that do not reveal unit information:
 - Tables where cells contain 3 or more contributing sample units; and
 - Details of fitted models, such as regression parameters.
- Examples of data that must be kept secure:
 - Prints of unit data;
 - Tabulations containing less than 3 contributing sample units; and
 - Details of unusual units that require special treatment in modelling.

6. Modes of access to data

The ABS makes unit data available by three different modes of access that give different levels of restriction and monitoring of user behaviour.



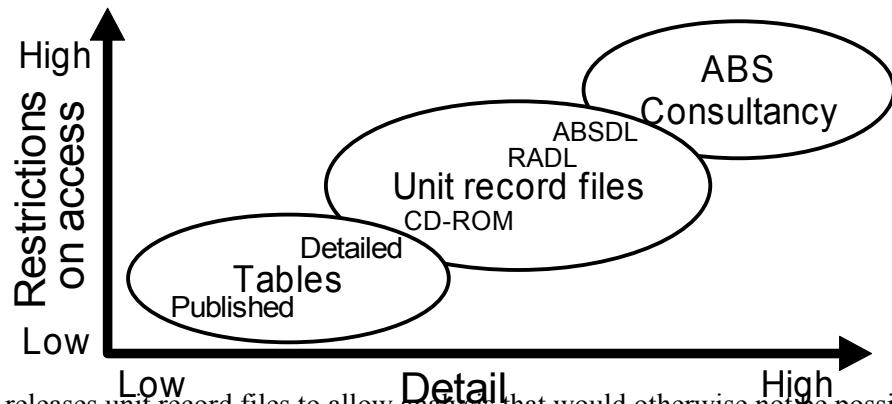
CD-ROM This allows you to access a unit record file stored within your organisation.

RADL Remote Access Data Laboratory allows you to submit programs and receive program outputs through a secure web site.

ABSDL ABS Site Data Laboratory allows you to do your analysis at the ABS on our equipment.

Unit record files form part of a spectrum of ABS products. As the level of detail required for your analysis increases, so do the access controls the ABS puts in place.

Modes of access to ABS information - level of access control by level of detail accessible



The ABS releases unit record files to allow analysis that would otherwise not be possible:

- CD-ROM access is for unit record files which of themselves have a low risk of identification because of the limited detail available.
- RADL access provides a number of extra protections against identification of individuals and is for files with a higher level of inherent risk.
- ABSDL access is for files with a particularly high inherent risk, perhaps because of the level of detail or because certain information on the file had been provided on behalf of respondents by an organisation and so could be easily matched by them.

Some analyses that require access to full detail of the data and the survey methodology on which it is based may best be performed by ABS staff through consultancy arrangements.

7. Accessing data on CD-ROM

CD-ROM access provides a unit record file to an organisation under conditions contained in an undertaking with the ABS. The organisation must keep track of its officers that have access to the unit information and ensure that all have signed an individual undertaking with the ABS.



Your Responsibility

Your access to unit record data is for statistical purposes within your organisation. If you change organisations you must surrender access, and ABS must be notified of the change.

Organisations receive a CD-ROM containing the unit data, and this must be kept securely. The unit data may be stored on computer systems in a secure manner that restricts access to **authorised individuals** - see section 5 for some guidelines on secure storage.

Copies of both organisational and individual application/undertakings are available on the ABS web site - see Appendix 1 for address. **Key requirements on the organisation** are:

- to ensure that access to information about individuals is restricted to officers of the organisation who have signed an individual undertaking with the ABS;
- to ensure that these users do not attempt to identify individuals;
- to ensure that users do not match the unit data to other files of unit data;
- to provide facilities for secure storage of computer files and printouts that may reveal unit data and for keeping track of these files and printouts;
- to provide a secure method of destruction for computer files and printouts; and
- to allow officers of the ABS (and/or their nominated representatives) access as necessary to audit compliance with these rules.

Individual researchers have corresponding requirements in the above areas: particularly

- to keep the unit information secure; and
- to not match the unit information to any other unit.

The ABS will contact your organisation annually to determine if access to this CURF is still required, and which officers should have access. You must keep track of any unit record information you have stored, so that it can be destroyed (shredded or securely deleted) when access is no longer required.

8. Accessing data in the Remote Access Data Laboratory

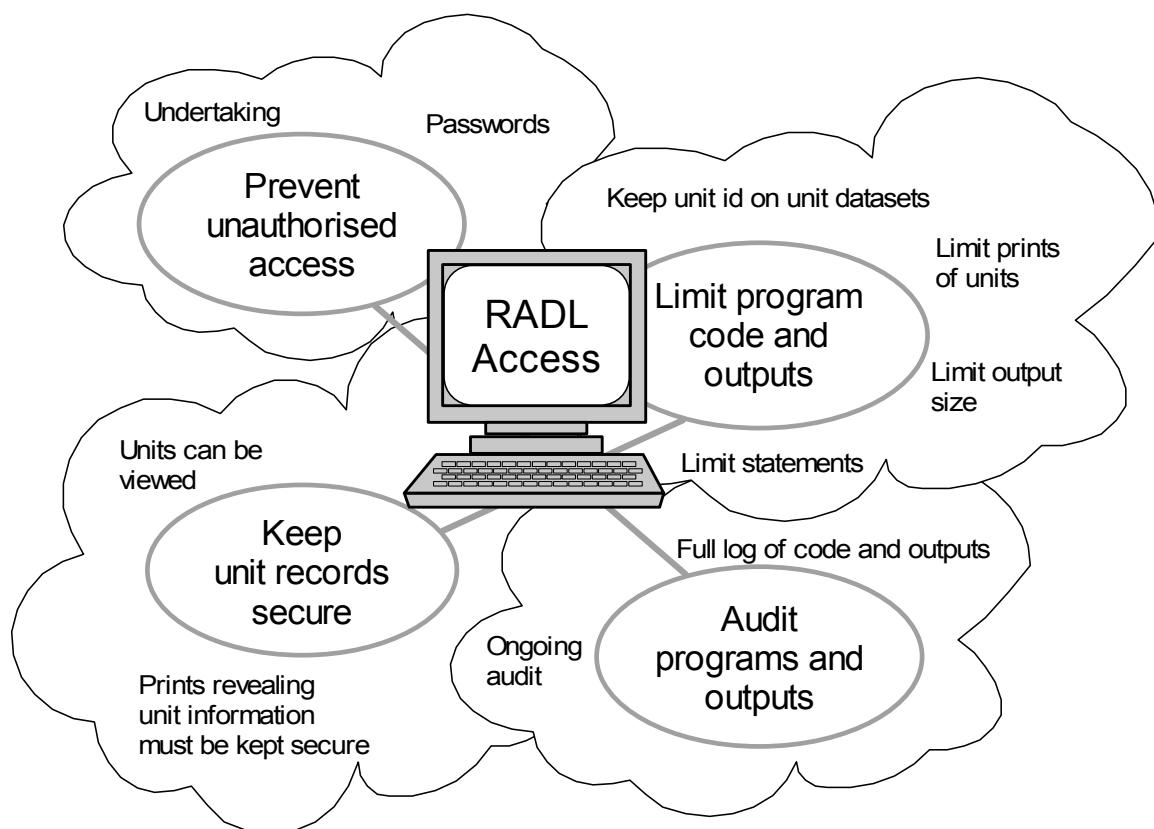
The RADL facility protects the unit record data by keeping the unit record files in the ABS. Users submit program code via a secure web site, and program outputs are returned promptly if they pass a number of automatic checks. Program outputs that fail these checks can be requested, subject to a manual check.



Your Responsibility

Read carefully and abide by the conditions of use in the RADL User Guide. Keep secure any program outputs that reveal unit information, and destroy them in a secure manner if they are no longer needed.

Four key areas of protection for unit record information in the RADL



RADL user responsibilities

Some key responsibilities of users of the RADL for each area of protection are highlighted below. For more details (e.g. specific limits on code, or names of ABS unit identifiers) see the **RADL User Guide** - available through the contacts given in Appendix 1.

Preventing unauthorised access

- Users receive a **password** for access to the RADL. Do not reveal your password to anyone - you will be held responsible for activity conducted using your password.

Keeping unit record information secure

- You will be allowed to see certain outputs that reveal unit record information. These include:
 - Prints of a limited number of items for a limited number of **unit records**;
 - Tables containing one or more **small cells** that may be small enough to contain data from a single contributor;
 - Any other output that the ABS tells you to store securely or delete;
- These outputs (and derived outputs that still reveal the unit information) must be stored in a **secure** manner. The ABS recommends you destroy them as you finish with them. See section 5 for some guidelines on secure storage and deletion.

Limiting program code and outputs

- You are required to ensure that any programs you submit to the RADL comply with certain limitations outlined in the RADL User Guide. In particular:
 - All datasets of unit records **must include** the **ABS unit identifiers**.
 - Certain **program statements** must not be used in the RADL.
- Prints of unit records and some tabular outputs are diverted to a separate output file, so that you can easily store them securely or destroy them after viewing.
 - Such prints will not be performed at all if they contain too many variables or records.
- Other output will be returned automatically if it passes checks of:
 - Tabular size, density and structure;
 - Size of prints; and
 - Size of total outputs.
- Outputs failing these checks will be retained by the ABS
 - Users can request the release of such outputs, subject to further checks by the ABS to ensure an acceptable level of unit record data is released.
- On request, and under special circumstances, users may obtain output that would not normally be released.
 - The program and its output will be checked.
 - Users must comply with any security recommendations.

Auditing programs and outputs

- Users are required to cooperate with any ABS audit directions in respect to their code and output.
 - The ABS maintains a log of code and outputs run in the RADL securely stored within the ABS computing environment to be used for administrative and auditing purposes only.

9. Accessing data in the ABS Data Laboratory

The ABSDL facility protects the unit record data by keeping the CURF physically in the ABS. The user accesses the data on an ABS computer, on site at the ABS.



Your Responsibility

Read carefully and abide by the conditions of use for the ABSDL. Do not bring any information in or out of the ABSDL except by the means provided by the ABS.

Security arrangements for the ABSDL

- Users must book ahead of time for each use of the ABSDL.
- Users must sign in and out of the ABSDL.
- Printouts and personal data transmitting devices must not be taken in or out of the ABSDL.

Getting information in and out of the ABSDL

- The ABS provides a means for users to bring information (including program code and outputs) in or out of the ABSDL in the form of electronic files.
 - All such files must be in formats that can be read within the ABS.
 - The user must describe the information on each file to assist with clearance.
 - Information may not enter or leave the ABSDL except via the ABSDL clearance officer.

Getting output cleared

- Output may leave the ABSDL only after it has been cleared by ABS staff.
 - The ABS maintains a full log of all information in the ABSDL.
 - User should provide explanation of required outputs, providing programs and logs where necessary.
 - Prints of unit records and output based on single contributors will not be released
 - Cleared output will be electronically returned to the user within 48 hours.

10. Taking your responsibilities seriously

The ABS is serious about the organisation and user responsibilities described in this manual. It is the trust that the Australian Statistician places in the user community that enables the ABS to continue to make available unit record information.



Your responsibility

If you breach the conditions of the undertaking, the ABS may respond by:

- Pointing out your error and issuing a warning;
- Removing access to the data from you and your organisation; or even
- Taking legal action.

The ABS will also take action against an organisation that has allowed its researchers to disregard the conditions of the undertaking

The undertakings that organisations and their individual researchers sign with the ABS provide a basis for legal action in cases where conditions of the undertaking have been wilfully ignored. The ABS also retains the right to cancel access to unit record data for organisations or specific individuals. The need to impose such sanctions is not expected to arise - but the ABS will not hesitate to act as required to ensure that conditions of the undertaking are being upheld.

11. Any questions?

If you are unclear about any issue raised in this manual, or have other questions about your responsibilities as a user of ABS CURF data, please contact us.



Contact information for the ABS CURF Management Unit is provided in appendix 1.

The ABS continually updates information about CURFs on the Access to ABS CURFs page on the ABS web site (www.abs.gov.au). From the ABS home page select Services We Provide, then Confidentialised Unit Record Files (CURFs).

Users should refer to this page regularly for the most up to date information and to access:

- Information about applying for, and renewing access to CURFs
- The RADL User Guide
- CURF Application and Undertaking forms
- A List of Available CURFs
- A List of CURFs Expected to be Released
- A List of CURF Responsible and Contact Officers
- Information about how CURFs have been used by researchers
- Information about the ABS/Australian Vice-Chancellors' Committee CURF Agreement
- CURF Prices

Appendix 1: Administrative arrangements for access to unit data

Applying for access to ABS confidentialised unit record files (CURFs)

Look on the
ABS web site
www.abs.gov.au

To apply for access

- to a given CURF
- for a given access mode:
 - CD-ROM
 - RADL
 - ABS_DL

Clients wishing to apply for access to ABS CURFs, should refer to the **Access to ABS CURFs** web page. From the ABS home page select Services We Provide, then Confidentialised Unit Record Files (CURFs).

An **organisation** must submit an Application and Undertaking form signed by the organisation's Responsible Officer for CURFs it requires to be available to its members.

An **individual** must submit an Application and Undertaking form, countersigned by his/her organisation's CURF Contact Officer to access CURFs he/she wishes to use.

- This is in addition to the forms for his/her organisation

Application and Undertaking forms are available from the Access to ABS CURFs web page.

Contact information

For further CURF information contact:

CURF Management Unit
Information Management & Census Division
Australian Bureau of Statistics

Email: curf.management@abs.gov.au
Ph: (02) 6252 5731 or (02) 6252 5853

Postal:
Locked Bag 10
Belconnen ACT 2616

Physical:
45 Benjamin Way
Belconnen ACT 2617

For further RADL information contact:

Email: radl@abs.gov.au

This information is effective 8 March 2005 and any subsequent revisions will be published on the ABS web site.